The "Tastes, Ties, and Time" (T3) research project has caused controversies, by collecting and using data of online Facebook network and offline on-campus information from the Class 2009 of a "diverse private college in the Northeastern U. S. ". Some researchers have been attracted to use this data (Lewis et. al., 2008; Wimmer & Lewis, 2010; Lewis et. al., 2012), while others cast concerns on the propriety of its existence (Stutzman, 2008; Hargittai, 2008; Zimmer, 2010). From my own point of view, collecting and using the T3 data might have violated two of Salganik's four principles of ethical research, while it managed to adhere to the other two. If given the access, I would use this data for my own research. Because the potential harm brought by its existence is fading away with the time. And scientific research with it, if using it properly, could be beneficial to both academia and public interests, and provide guidance to future data collection, such as direction or more thorough ethnic concerns. But still, researchers need to be careful when collecting and using data in the future, making every effort to adhere to Salganik's four principles of ethnical research.

Although the dataset is now unavailable, we could still get to know its main content from papers relevant to either research or the debate. The greatest benefit and potential harm of it are basically the same: it not only contains online network data of these student, but also contains some information of their real life. To social network analysis, it could be a great treasure, for it provides the evolution of a concrete network. But problems occur because it seriously divulged personal information of the Class 2009 student without their consent. Although the data collectors have made great effort to anonymize the data, the source of it was still easily tracked.

These might have violated some of the four principles of ethical research. According to Salganik, they are: (1) respect for persons, (2) beneficence, (3) justice, and (4) respect for law and public interest. The collection and use of the T3 data might have violated principle (1) and (4), and might be adherence to the other two principles.

The first principle, respect for persons, requires researchers to treat individuals as autonomous (Salganik, in open review). This means it should have asked for students' consent. But the researchers have not taken this step in data collection. The data was collected "with permission from Facebook and the university in question" (Lewis, 2008), which means the collection of this data was not known by the subjects when it was formed. And undoubtedly, offline information, including accommodation and major of the students, was collected with authorization of the university, which happened on the institution level. No effort has been shown to have made to ask for students' own consent. Furthermore, without the process of asking for authorization from the students, the T3 data was released to researchers outside of the research group collecting it. Although a Terms and Conditions of Use statement was set as a necessary procedure to get access to the data, it was actually more for the interest of the research group, instead of the students themselves.

The principle (4) might have also been violated. The transparency-based accountability, required by this principle, indicates that the researchers should have been clear about the results of their research process at all stage and take the responsibility of their action (Salganik, in open review). However, the T3 data was firstly released, and then made unavailable online. This means that the researchers had not got good knowledge of possible results of releasing the data when they did that.

As far as I think, the T3 project does not necessarily violate the second and third principle.The second principle, beneficence, requires the researchers to do their research without harming the subjects, or at least to maximize possible benefits meanwhile minimizing possible harms (Salganik, in open review). Although the researcher failed to collect data without harming the students' privacy, they have done their best to protect them by anonymizing information indicating subjects' identity. The problem is they were too optimistic to predict that the identify of the university and its students would eventually get divulged.

And for the third principle, justice, it requires researchers to avoiding negatively influence the vulnerable groups to satisfy benefits of the privileged (Salganik, in open review). In the case of T3 project, that's not the case. Although it may be argued that the researchers have power over the student, by getting access to their profiles and online network beyond their consent, the autonomous students are not actually vulnerable/disadvantaged group. They should have been protected by Facebook's terms of services and their university. Thus actually, they are Facebook and the university that should take the responsibility of this part.

If I could have access to the T3 data, I would use it for my own research. First, years has passed, and the data is no longer recent information of the subjects, which would control the potential harm to their security and privacy to a very limited stage. More importantly, Unlike the biased and incomplete data collected through traditional surveys, the T3 data or similar dataset could provide a relatively unbiased and concrete picture of online social network. And being longitudinal data, it could help to reveal the formation and evolution of social network. More excitingly, the data combines online and offline information, which would help to analyze the

interaction between social network and other factors, contributing to discussions such as racial segregation. Using this data for academic research is beneficial to understanding of the dramatically growing online world, which is beneficial to not only academic research but also the society, providing the foundation of knowledge to improve social integration. Second, diving into big or new data could also provide experience for future data collection and the development of research ethnic. Certainly, we should not neglect that in some cases, collected data, such as German census data of the Jewish, for the purpose of public interest and scientific research may cause unintended negative consequences. However, with the development of data collection and data mining, data could still be gathered anyway, even if it is for some unjust purposes. It is the purpose of using that has caused problems, not the datasets themselves. It's the

In conclusion, the T3 project might have violated some of Salganik's principles for ethnic research while is still adherence to the others. I would like to use the T3 data or other similar dataset for the reasons as above.

References:

Lewis, Kevin, Jason Kaufman, Marco Gonzalez, Andreas Wimmer, and Nicholas Christakis. "Tastes, ties, and time: A new social network dataset using Facebook. com." *Social networks* 30, no. 4 (2008): 330-342.

Lewis, Kevin, Marco Gonzalez, and Jason Kaufman. "Social selection and peer influence in an online social network." *Proceedings of the National Academy of Sciences* 109, no. 1 (2012): 68-72.

Nissenbaum, Helen. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press, 2009.

Narayanan, Arvind, and Vitaly Shmatikov. "De-anonymizing social networks." In *2009 30th IEEE symposium on security and privacy*, pp. 173-187. IEEE, 2009.

Salganik, Matthew J. *Bit by Bit: Social Research in the Digital Age*, Princeton University Press, Open review edition.

Wimmer, Andreas, and Kevin Lewis. "Beyond and below racial homophily: Erg models of a friendship network documented on facebook1." *American Journal of Sociology* 116, no. 2 (2010): 583-642.

Zimmer, Michael. ""But the data is already public": on the ethics of research in Facebook." *Ethics and information technology* 12, no. 4 (2010): 313-325.